An introduction to time series and time series models

Applied Time Series Analysis for Ecologists

Stockholm, Sweden — 24-28 March 2014

Topics for this morning

- Characteristics of time series (ts)
  - What is a ts?
  - Classifying ts
  - Trends
  - Seasonality (periodicity)

- Formal descriptions of ts
  - Expectation, mean & variance
  - Covariance & correlation
  - Stationarity

Topics for this morning

- Time series models
  - Autocovariance & Autocorrelation functions
  - Correlograms
  - White noise
  - Random walks
  - Autoregressive models
  - Moving average models
  - Autoregressive moving average models

What is a time series?

- A *time series* (ts) is a set of observations taken sequentially in time

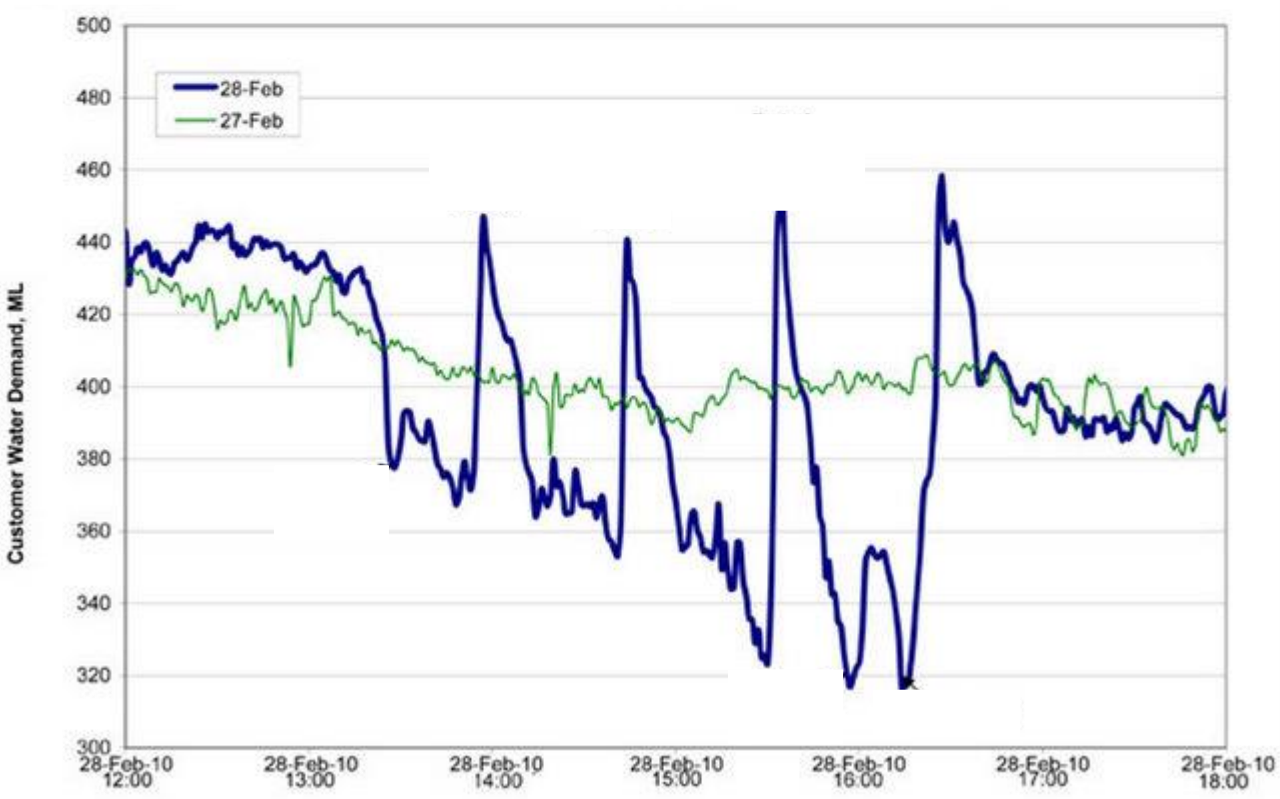
- A ts can be represented as a set

  $\{x_t : t = 1,2,3,\ldots,n\} = \{x_1,x_2,x_3,\ldots,x_n\}$

- For example,

  $\{10,31,27,42,53,15\}$

Example of a time series



Classification of time series (I)

I. By some index set

   A. Interval across real time $x(t)$; $t \in [1.1, 2.5]$

   B. Discrete time x_t

      1. Equally spaced; $t = \{1,2,3,4,5\}$
      2. Equally spaced w/ missing values; $t = \{1,2,4,5,6\}$
      3. Unequally spaced; $t = \{2,3,4,6,9\}$

Classification of time series (II)

II. By underlying process

   A. Discrete (eg, adults counted per minute)

   B. Continuous (eg, salinity, temperature)

Classification of time series (III)

III. By number of values recorded

   A. Univariate/scalar (eg, total # of fish caught)

   B. Multivariate/vector (eg, # of each spp of fish caught)

Classification of time series (IV)

IV. By type of values recorded

A. Integer (eg, fish caught in 5 min trawl = 2413)

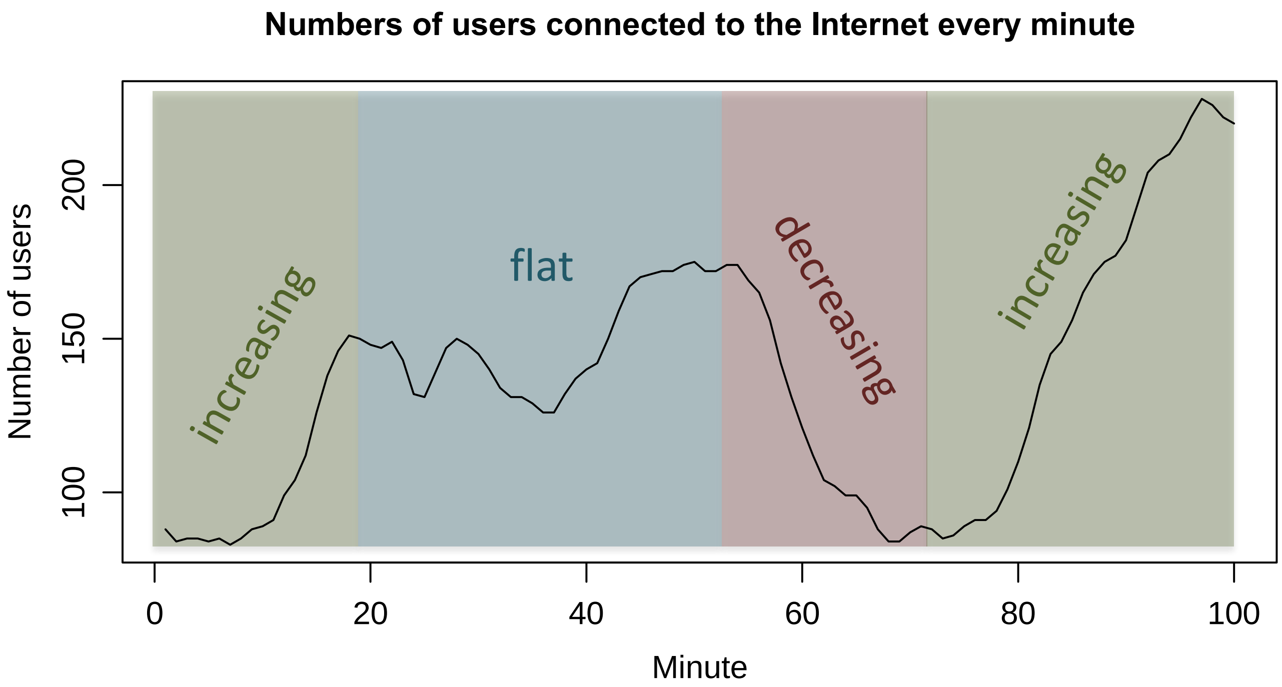
B. Rational (eg, fraction of marked birds = 47/951)

C. Real (eg, body mass = 10.2 g)

Statistical analyses of time series

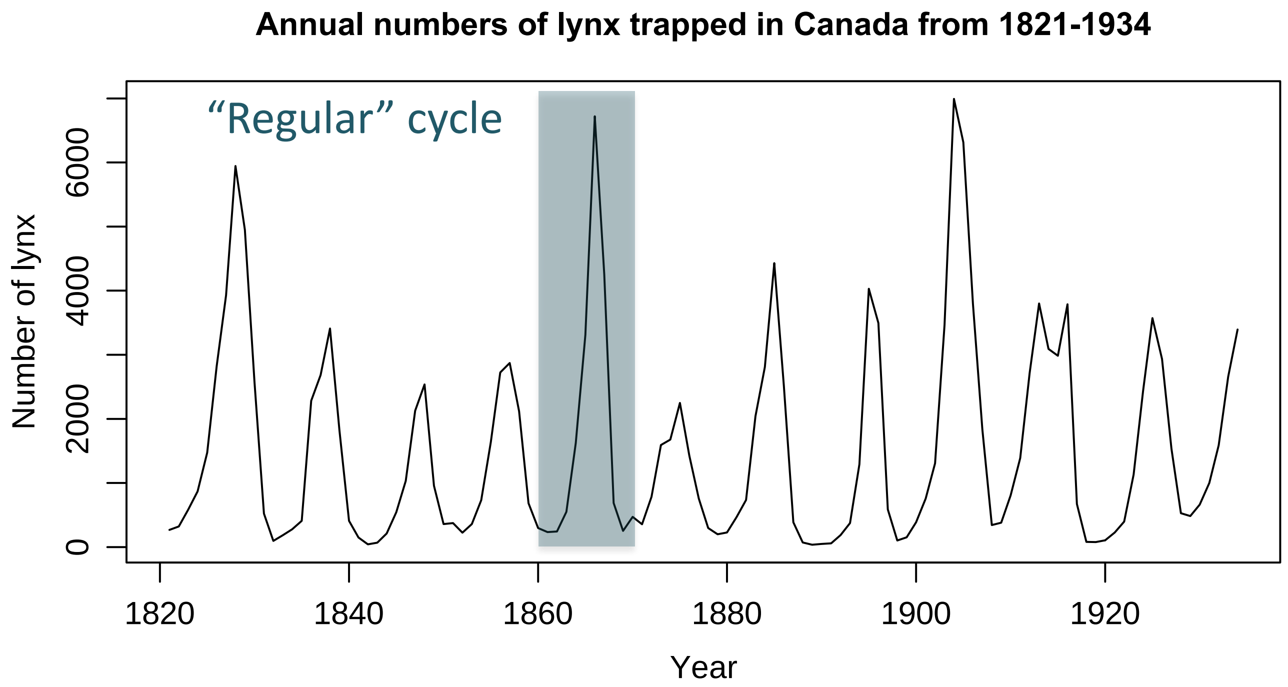
- Most statistical analyses are concerned with estimating properties of a population from a sample
- Time series analysis, however, presents a different situation
- Although one could vary the length of an observed sample, it is often impossible to make multiple observations at a given time (*eg*, one cannot observe today's exchange rate of SEK to USD more than once)
- This makes conventional statistical procedures, based on large sample estimates, inappropriate

Examples of time series



How would we describe this ts?

Examples of time series



How would we describe this ts?

What is a time series model?

- A *time series model* for $\{x_t\}$ is a specification of the joint distributions of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is thought to be a realization.

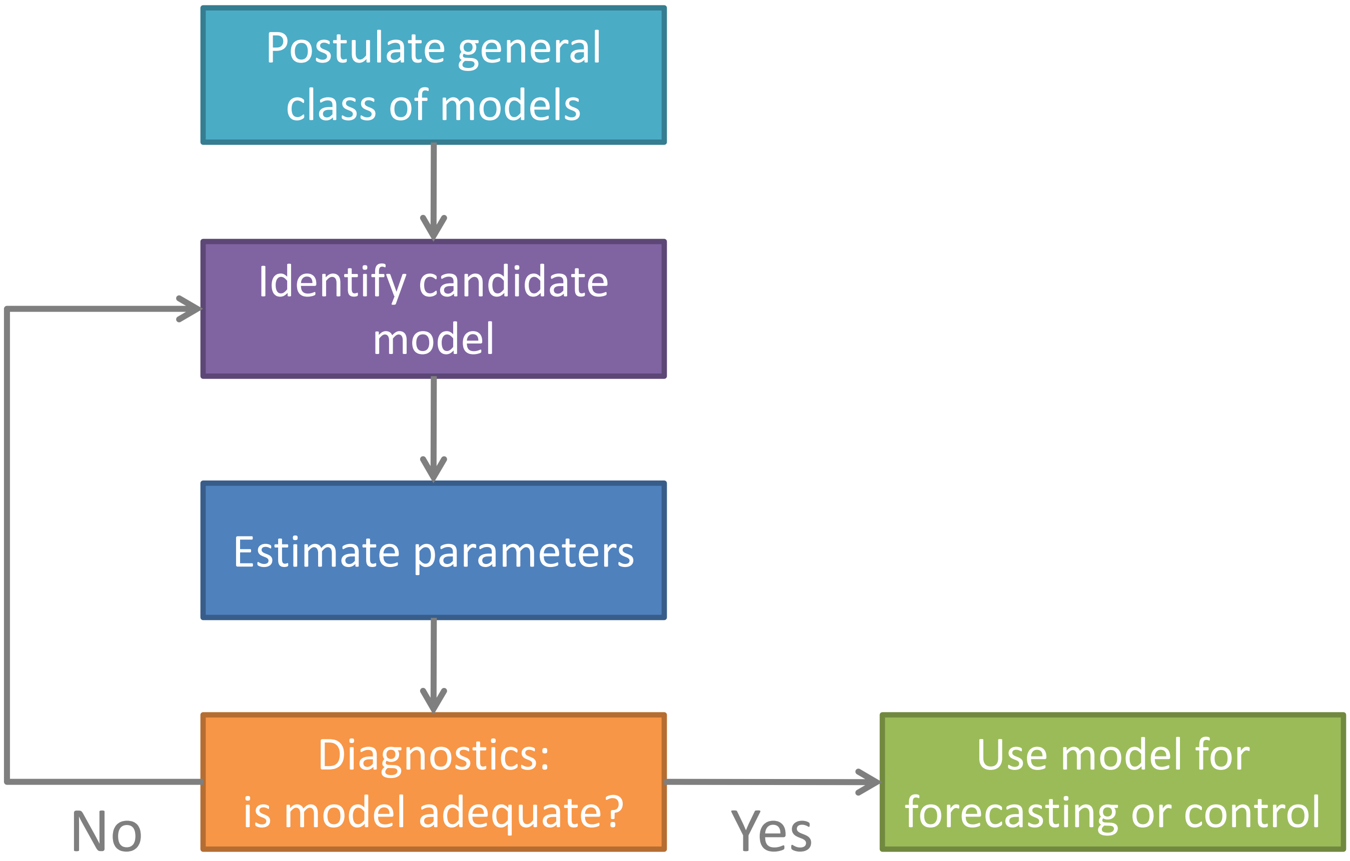
- For example,

  "white noise": $x_t = w_t$ and $w_t \sim N(0,0.1)$

  autoregressive: $x_t = 0.8x_{t-1} + w_t$ and $w_t \sim N(0,0.1)$

Iterative approach to model building

Also known as the "Box-Jenkins Approach"



Classical decomposition of time series

- *Classical decomposition* of an observed time series is a fundamental approach in time series analysis
- The idea is to decompose a time series $\{x_t\}$ into a trend (m_t), a seasonal component (s_t), and a remainder (e_t)

$$x_t = m_t + s_t + e_t$$

Expectation, mean & variance

- The *expectation* (E) of a variable is its mean value in the population

- $E(x) \equiv$ mean of $x = \mu$

- $E([x - \mu]^2) \equiv$ mean of squared deviations about μ

  $\equiv$ variance $= \sigma^2$

- Can estimate σ^2 from sample as

$$\mathrm{Var}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Covariance

- If we have 2 variables (x, y) we can generalize variance

$$\sigma_x^2 = \mathrm{E}\left[(x-\mu_x)(x-\mu_x)\right]$$

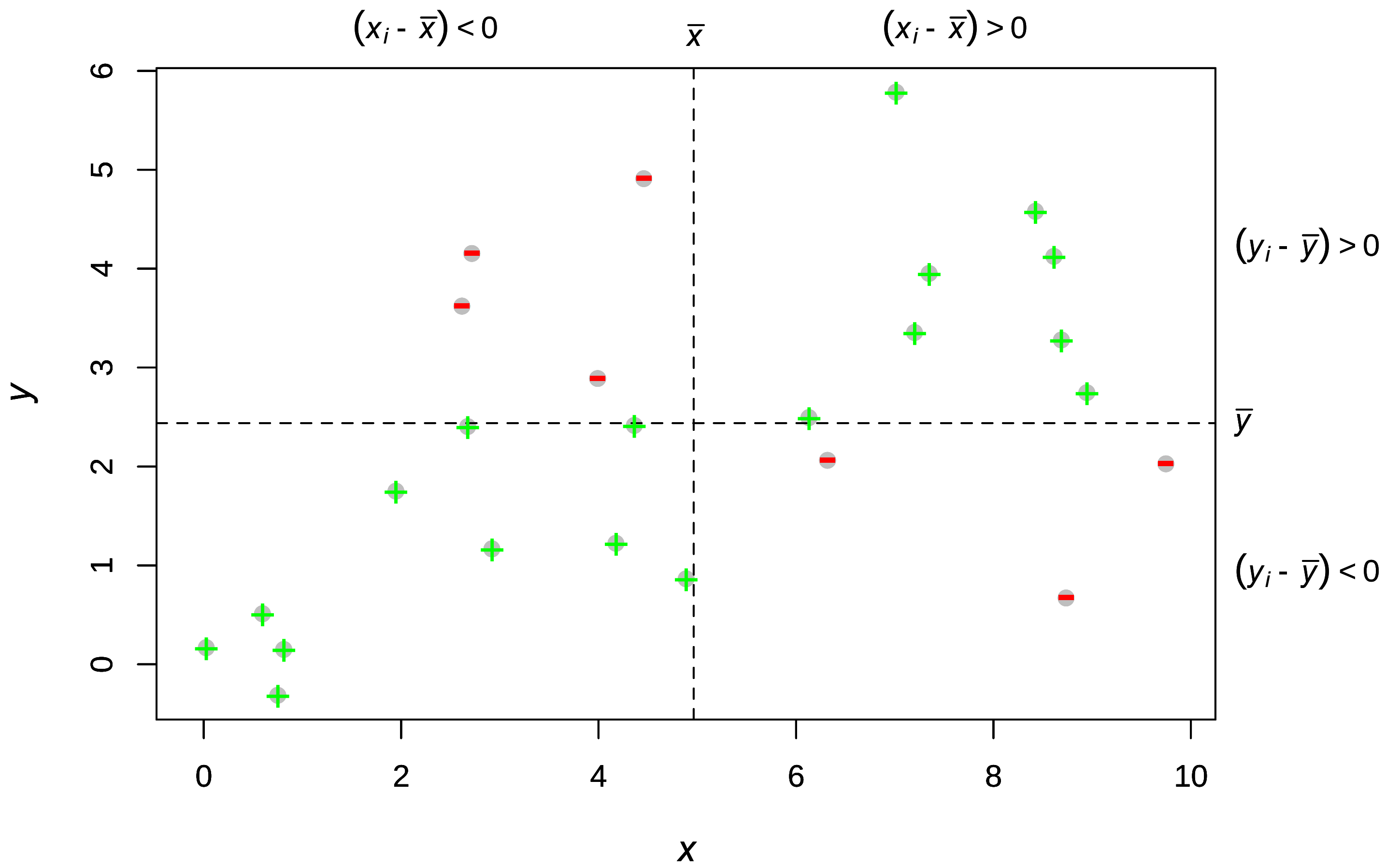
to *covariance*

$$\gamma(x,y) = \mathrm{E}\left[(x-\mu_x)(y-\mu_y)\right]$$

- Can estimate γ from sample as

$$\mathrm{Cov}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

Graphical example of covariance



Correlation

- *Correlation* is a dimensionless measure of the linear association between 2 variables *x* & *y*
- It is simply the covariance standardized by the standard deviations

$$\rho(x,y) = \frac{E\left[(x-\mu_x)(y-\mu_y)\right]}{\sigma_x \sigma_y} = \frac{\gamma(x,y)}{\sigma_x \sigma_y} \in [-1,1]$$

- Can estimate γ from sample as

$$\mathrm{Cor}(x,y) = \frac{\mathrm{Cov}(x,y)}{\mathrm{sd}(x)\mathrm{sd}(y)}$$

The ensemble & stationarity

- Consider again the mean function for a time series:

  $\mu(t) = E(x_t)$

- The expectation is taken across an *ensemble* (population) of all possible time series

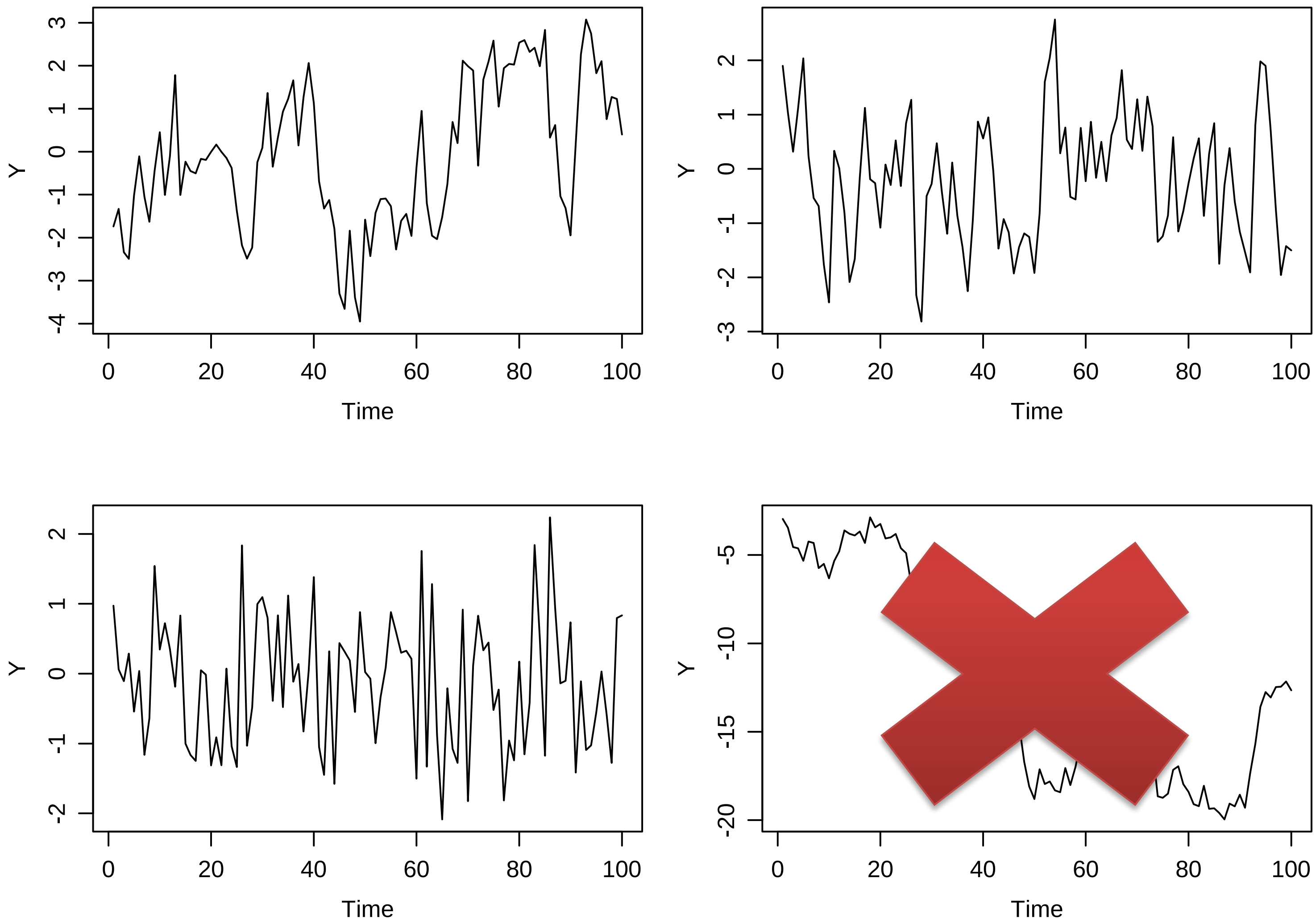
- With only 1 sample, however, we must estimate the mean at each time point by the observation

- If $E(x_t)$ is constant across time, we say the time series is *stationary* in the mean

Stationarity of time series

- *Stationarity* is a convenient assumption that allows us to describe the statistical properties of a time series.
- In general, a time series is said to be stationary if there is
  1) no systematic change in mean or variance,
  2) no systematic trend, and
  3) no periodic variations or seasonality

Which of these are stationary?



Autocovariance function (ACVF)

- For stationary ts, we can define the *autocovariance function* (ACVF) as a function of the time lag (k)

$$\gamma_k = \mathrm{E}\left[(x_t - \mu_x)(x_{t+k} - \mu_x)\right]$$

- Very "smooth" series have large ACVF for large k; "choppy" series have ACVF near 0 for small k

- Can estimate γ_k from sample as

$$c_k = \frac{1}{\boxed{n}} \sum_{t=1}^{\boxed{n-k}} (x_t - \bar{x})(x_{t+k} - \bar{x})$$

Autocorrelation function (ACF)

- The *autocorrelation function* (ACF) is simply the ACVF normalized by the variance

$$r_k = \frac{g_k}{s^2} = \frac{g_k}{g_0}$$

- ACF measures the correlation of a time series against a time-shifted version of itself (& hence “auto”)

- Can estimate r_k from sample as

$$r_k = \frac{c_k}{c_0}$$

Properties of the ACF

The ACF has several important properties, including

1) $-1 \leq r_k \leq 1$,
2) $r_k = r_{-k}$ (ie, it's an "even function"),
3) r_k of periodic function is itself periodic
4) r_k for sum of 2 indep vars is sum of r_k for each

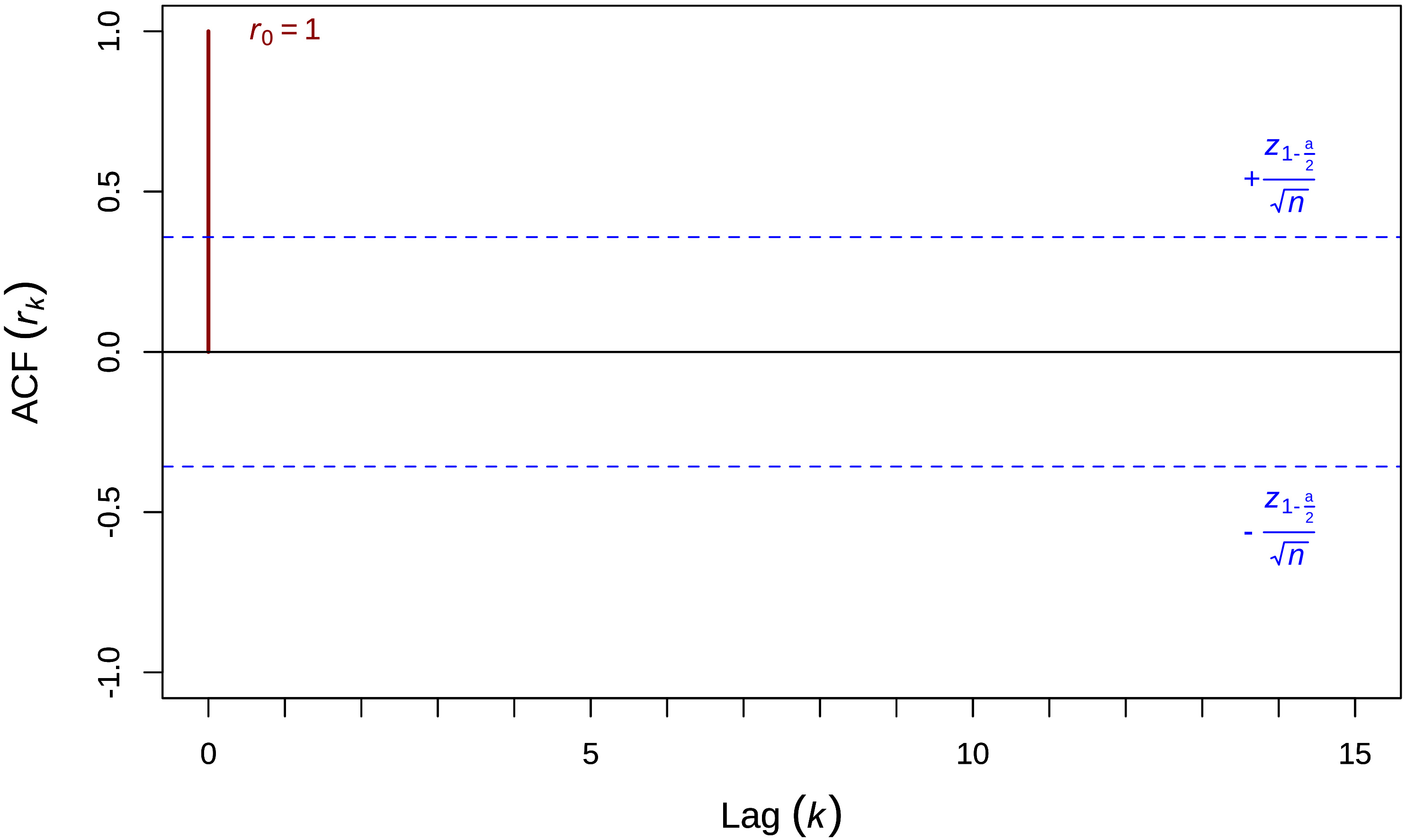
The correlogram

- The common graphical output for the ACF is called the *correlogram*, and it has the following features:

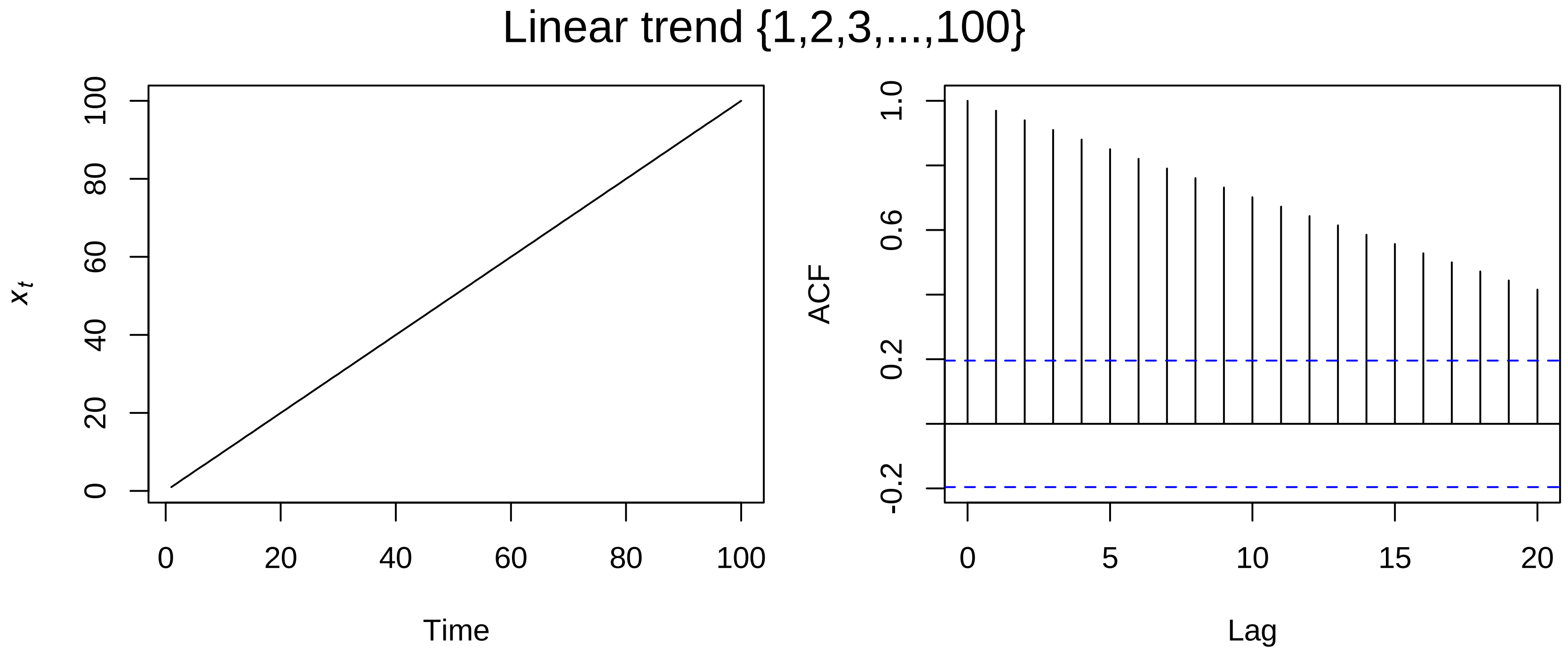
1) x-axis indicates lag (0 to k);
2) y-axis is autocorrelation r_k (-1 to 1);
3) lag-0 correlation (r_0) is always 1 (it's a ref point);
4) If ρ_k = 0, then sampling distribution of r_k is approx. normal, with var = 1/n;
5) Thus, a 95% conf interval is given by

$$\pm \frac{z_{1-a/2}}{\sqrt{n}}$$

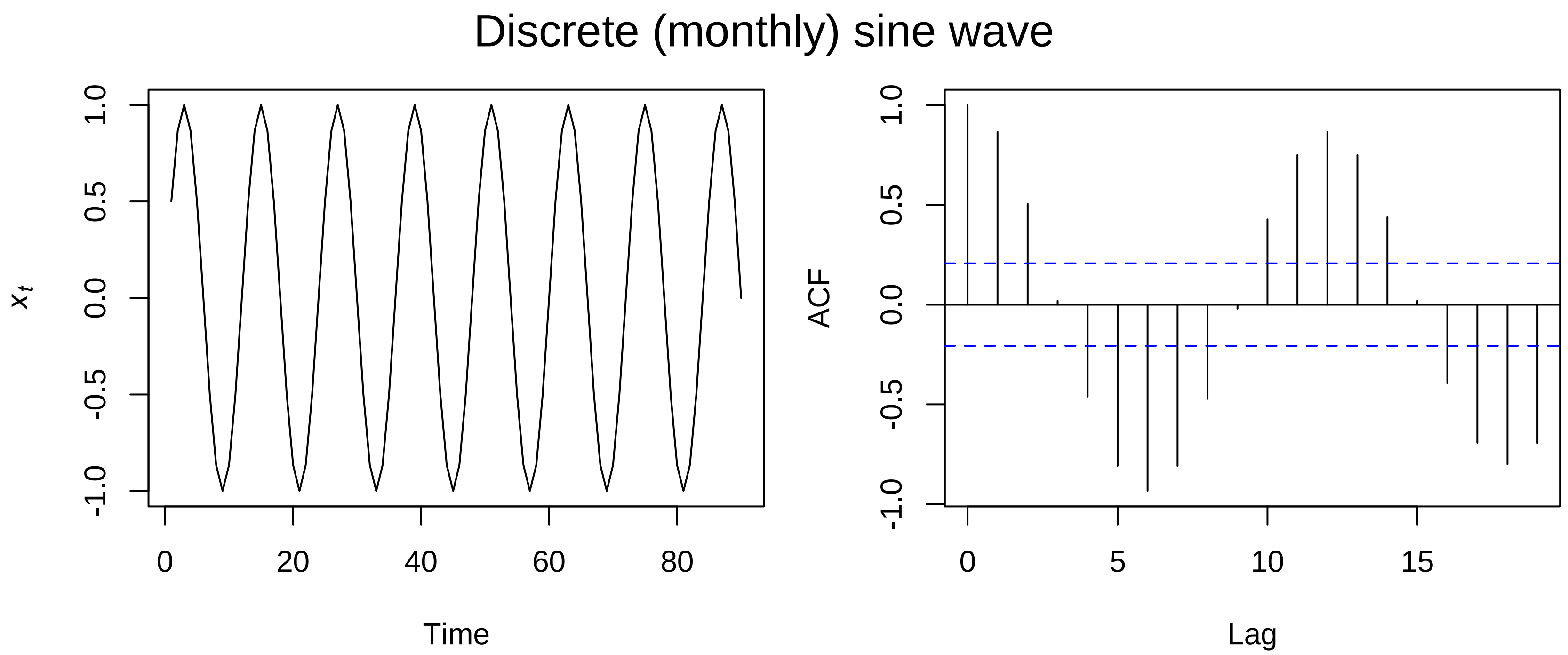
The correlogram



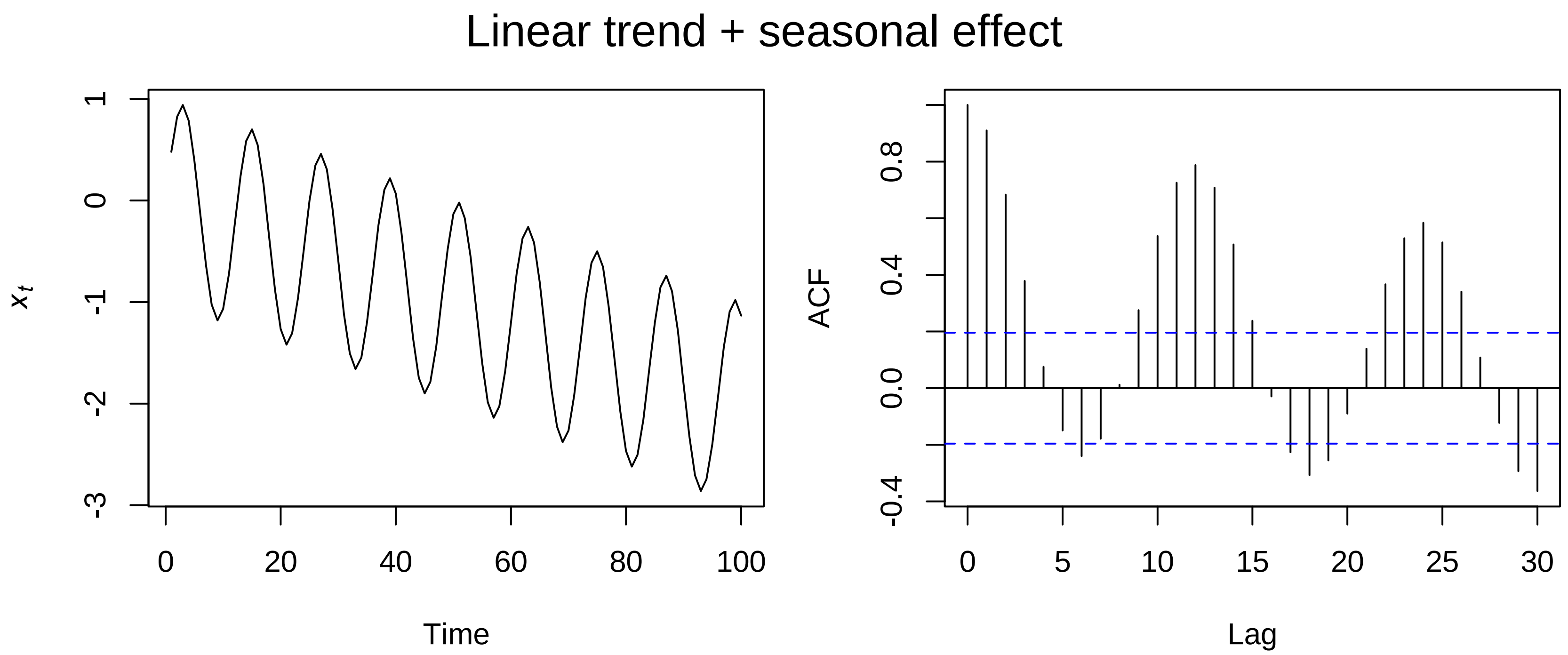
Correlogram for deterministic trend



Correlogram for sine wave

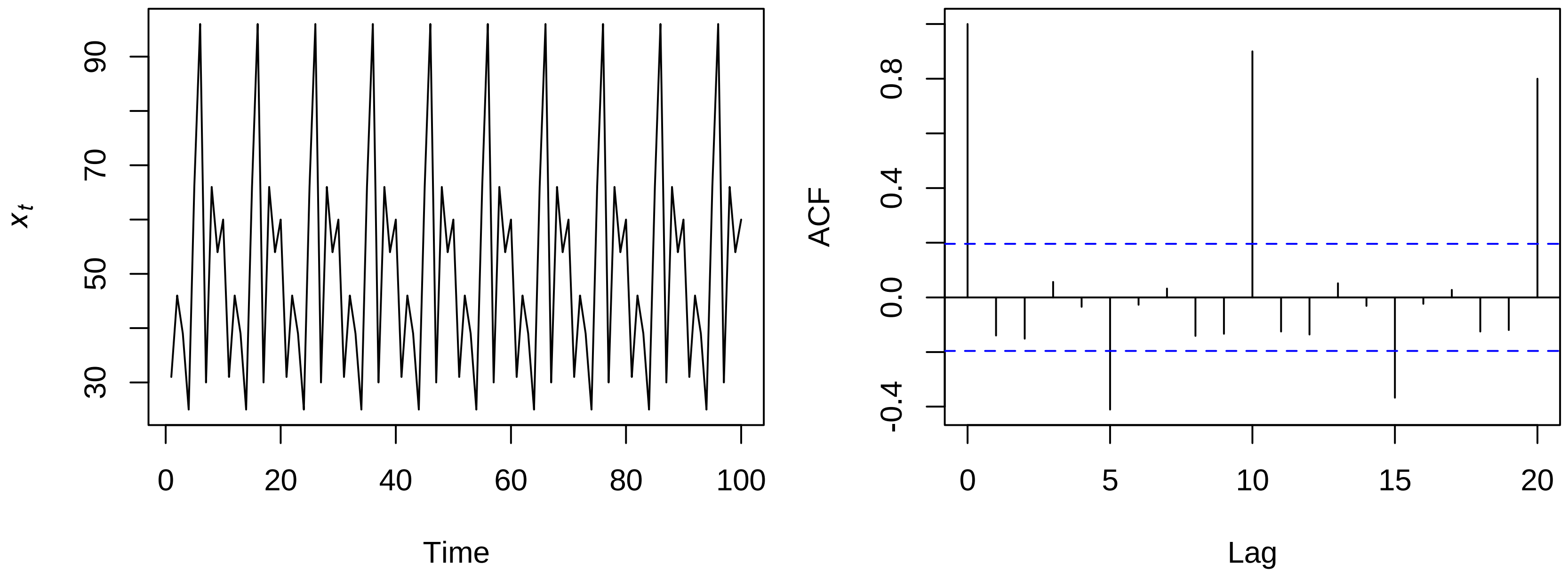


Correlogram for trend + season

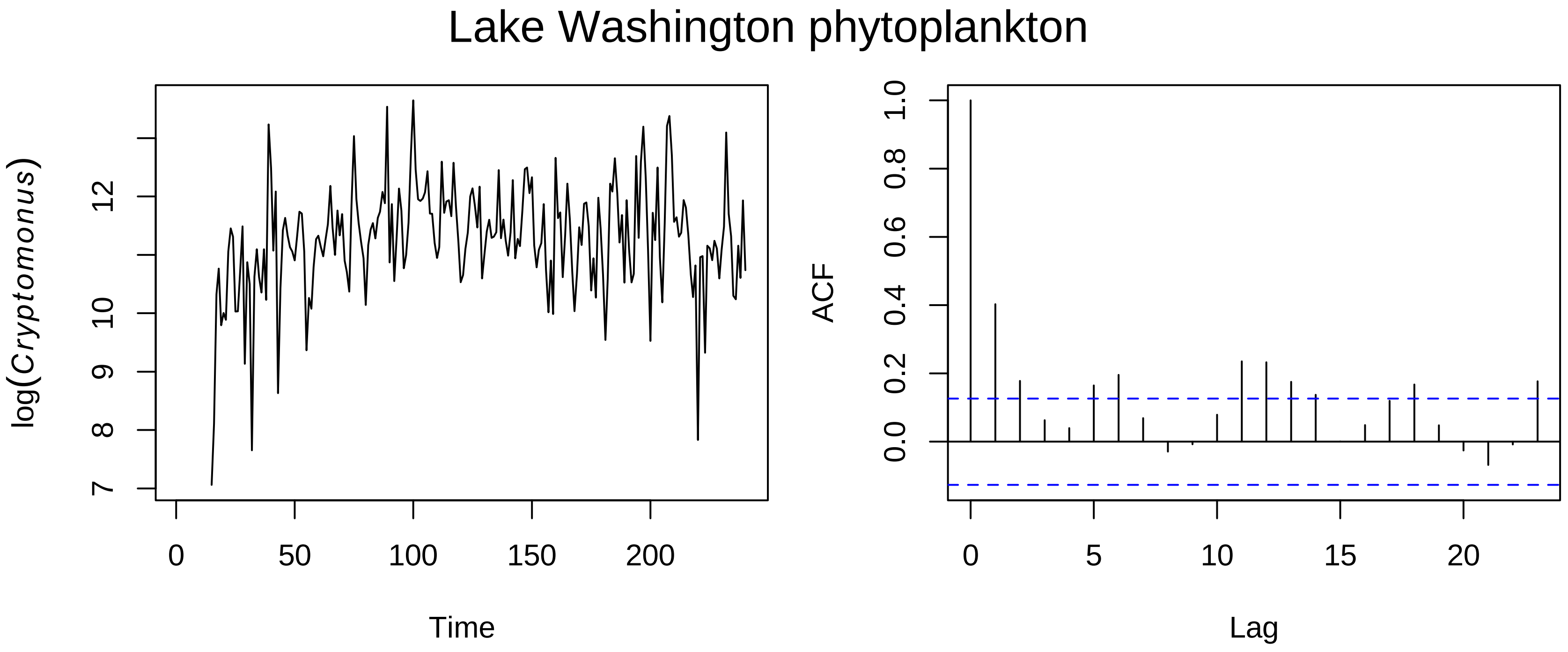


Correlogram for random sequence

Random sequence of 10 numbers repeated 10 times



Correlogram for real data



White noise (WN)

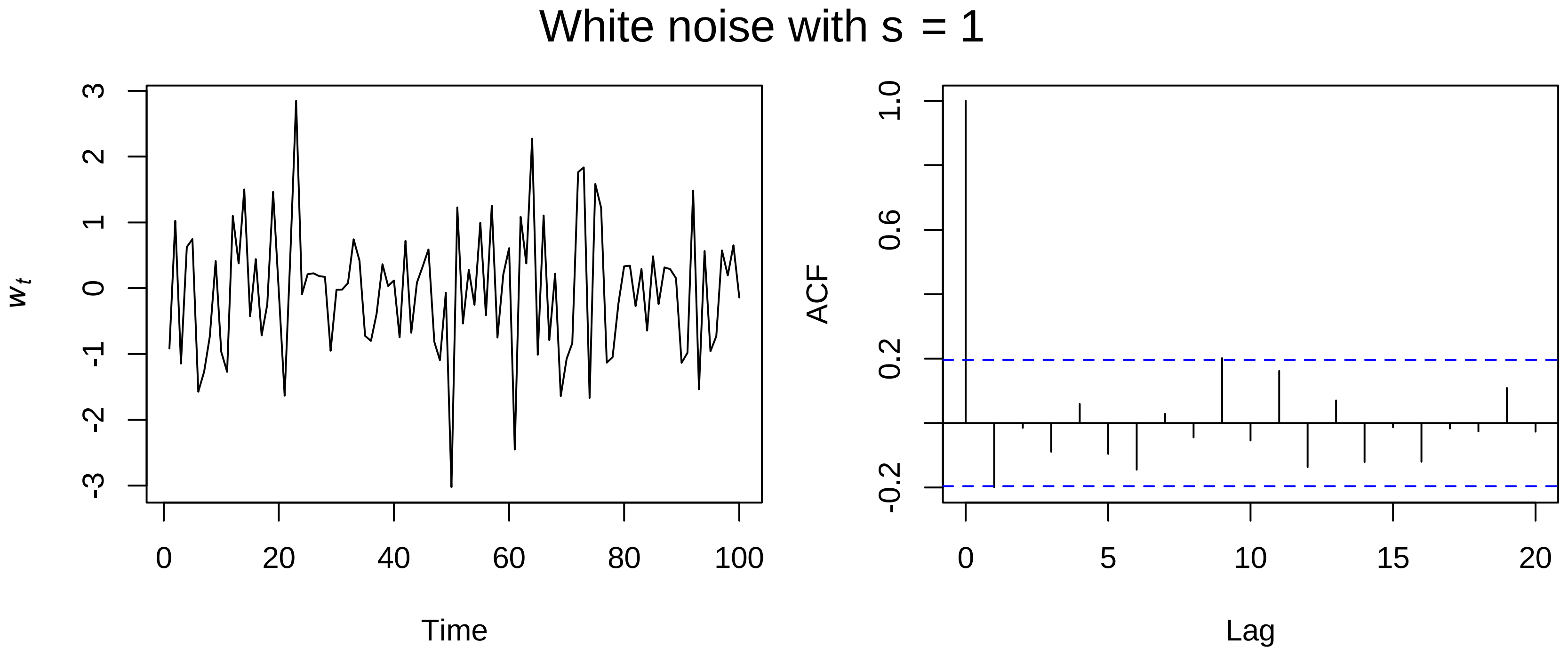
A time series $\{w_t : t = 1,2,3,\ldots,n\}$ is *discrete white noise* if the variables $w_1, w_2, w_3, \ldots, w_n$ are

1) *independent*, and
2) *identically distributed* with a mean of zero

WN has the following 2nd-order properties:

$$\mu_w = 0 \qquad \gamma_k = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \qquad \rho_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$$

White noise


White noise with s = 1
w_t
3
2
1
0
-1
-2
-3
0
20
40
60
80
100
Time
ACF
1.0
0.6
0.2
-0.2
0
5
10
15
20
Lag

Random walk (RW)

A time series $\{x_t : t = 1,2,3,\ldots,n\}$ is a *random walk* if

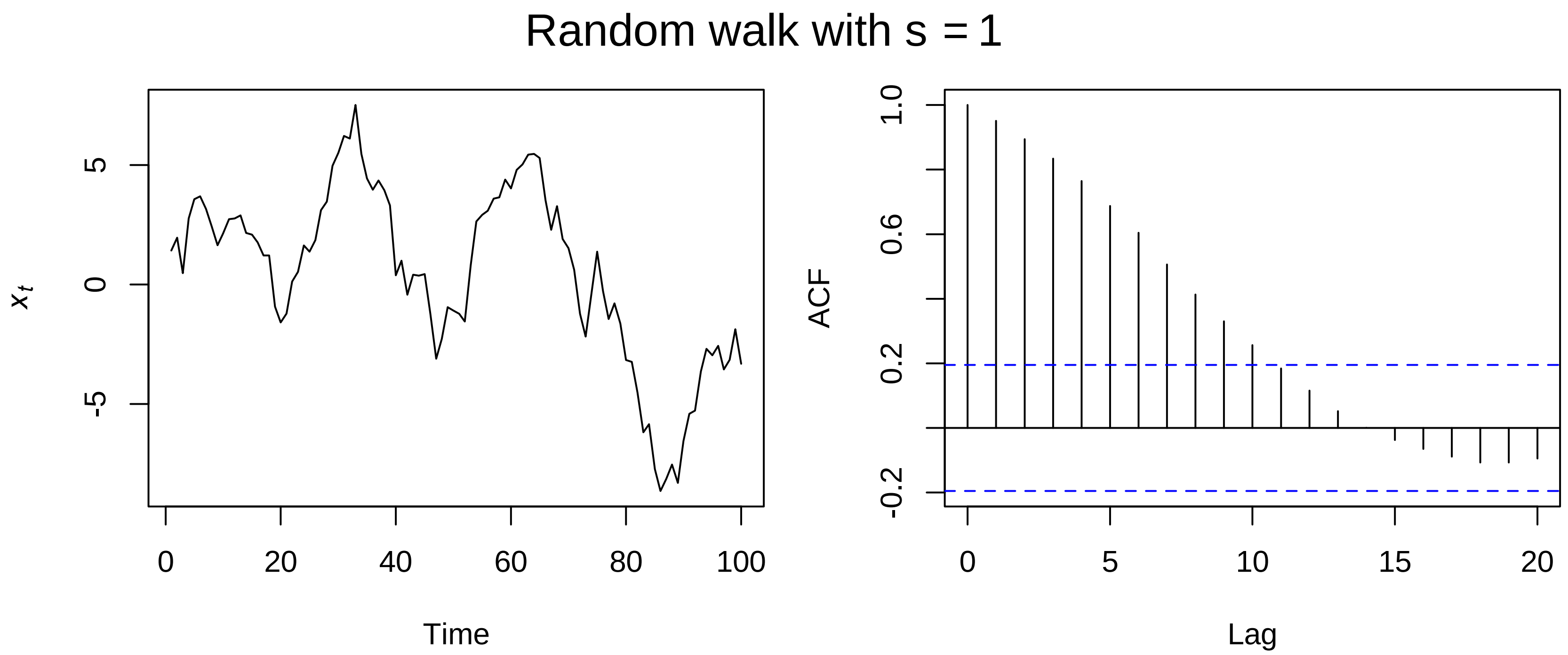
1) $x_t = x_{t-1} + w_t$, and
2) w_t is white noise

RW has the following 2nd-order properties:

$$\mu_w = 0 \qquad \gamma_k(t) = t\sigma^2 \qquad \rho_k(t) = \frac{t\sigma^2}{\sqrt{t\sigma^2 (t+k)\sigma^2}} = \frac{1}{\sqrt{1 + k/t}}$$

Random walks are NOT stationary!

Random walk (RW)



The difference operator (∇)

- Define the first *difference operator* as

$$\nabla x_t = x_t - x_{t-1}$$

- Differences of order *d* are then defined by

$$\nabla^d = (1 - B)^d$$

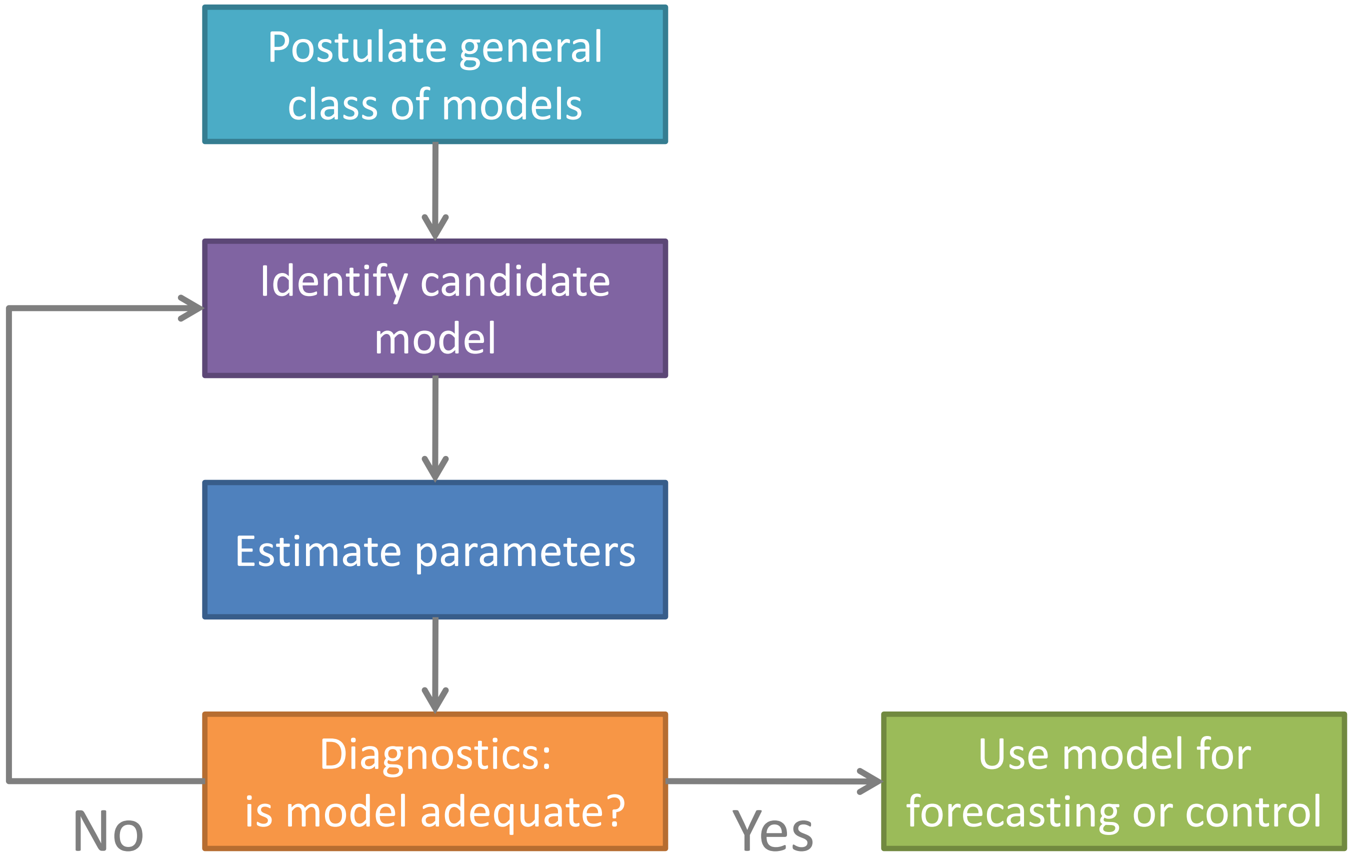
- So, first differencing a RW model yields WN

$$x_t - x_{t-1} = w_t$$

Difference to remove trend/season

- Differencing is a very simple means for removing a trend or seasonal effect
- The 1st-difference removes a linear trend, a 2nd-difference would remove a quadratic trend, etc.
- For seasonal data, using a 1st-difference with *lag = period* removes both trend & seasonal effects
- Pro: no parameters to estimate
- Con: no estimate of stationary process

Iterative approach to model building



Linear stationary models

- Linear filters are a useful way of modeling time series
- Here we extend those ideas to a general class of models call autoregressive moving average (ARMA)

Autoregressive (AR) models

- An *autoregressive* model of order *p*, or AR(*p*), is defined as

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t$$

  where we assume

  1) w_t is WN, and
  2) $\phi_p \neq 0$ for order-*p* process

- *Note*: RW model is special case of AR(1) with $\phi_1 = 1$

The backward shift operator (**B**)

- Define the *backward shift operator* by

$$\mathbf{B}x_t = x_{t-1}$$

- Or, more generally as

$$\mathbf{B}^k x_t = x_{t-k}$$

- So, RW model can be expressed as

$$x_t = \mathbf{B}x_t + w_t$$

$$(1-\mathbf{B})x_t = w_t$$

$$x_t = (1-\mathbf{B})^{-1} w_t$$

Stationary AR models

- We can write out an AR(*p*) model using the backward shift notation, such that

$$\phi_p(\mathbf{B})x_t = \left(1 - \phi_1\mathbf{B} - \phi_2\mathbf{B}^2 - \ldots - \phi_p\mathbf{B}^p\right)x_t = w_t$$

- If we treat **B** as a number, we can write the *characteristic equation* as

$$\phi_p(\mathbf{B}) = 0$$

- In order to be stationary, *all* roots of characteristic equation must exceed 1 in absolute value!

Examples of (non)stationary models

- Consider this simple AR(1) model:

  $x_t = \phi x_{t-1} + w_t$

  $(1 - \phi\mathbf{B})x_t = w_t$

- A random walk $x_t = x_{t-1} + w_t$ is not stationary because

  $\phi = 1$

  $\phi(\mathbf{B}) = 1 – \mathbf{B} = 0$ and hence $\mathbf{B} = 1$

- However, the AR(1) model $x_t = 0.5x_{t-1} + w_t$ is because

  $\phi = 0.5$

  $\phi(\mathbf{B}) = 1 – 0.5\mathbf{B} = 0$, and hence $\mathbf{B} = 2$

Examples of AR(1) processes

Examples of AR(1) processes

Partial autocorrelation function

- The partial *autocorrelation function* (PACF) measures the linear correlation of a series x_t and x_{t+k} with the linear dependence of $\{x_{t-1},x_{t-2},\ldots,x_{t-(k-1)}\}$ removed
- It is defined as

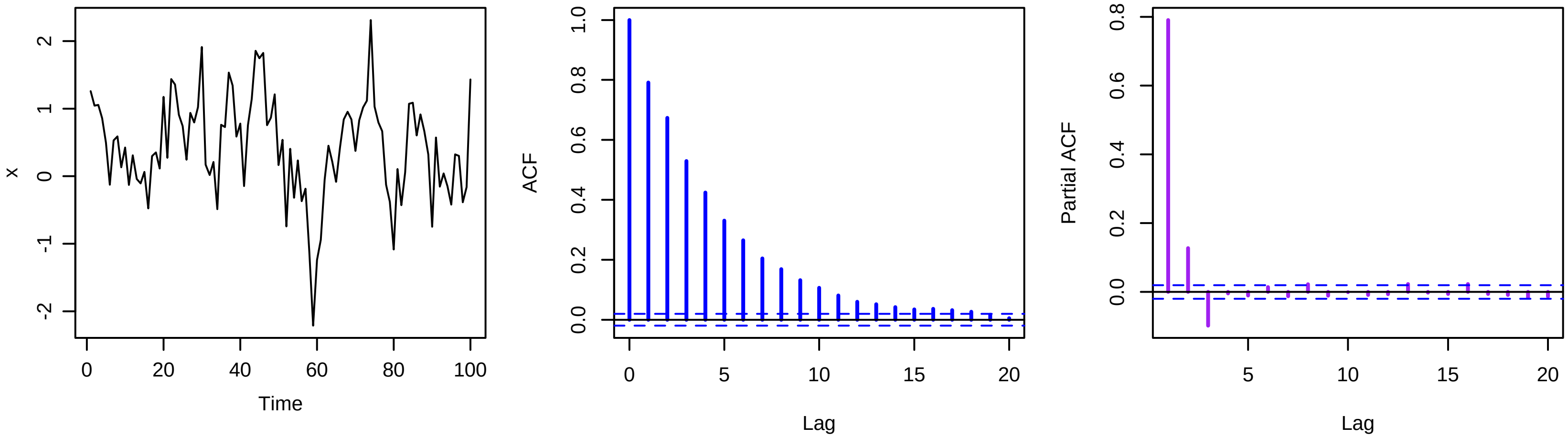
$$-1 \leq \phi_{kk} \leq 1$$

$$x_k^{k-1} = \beta_1 x_{k-1} + \beta_2 x_{k-2} + \cdots + \beta_{k-1} x_1$$

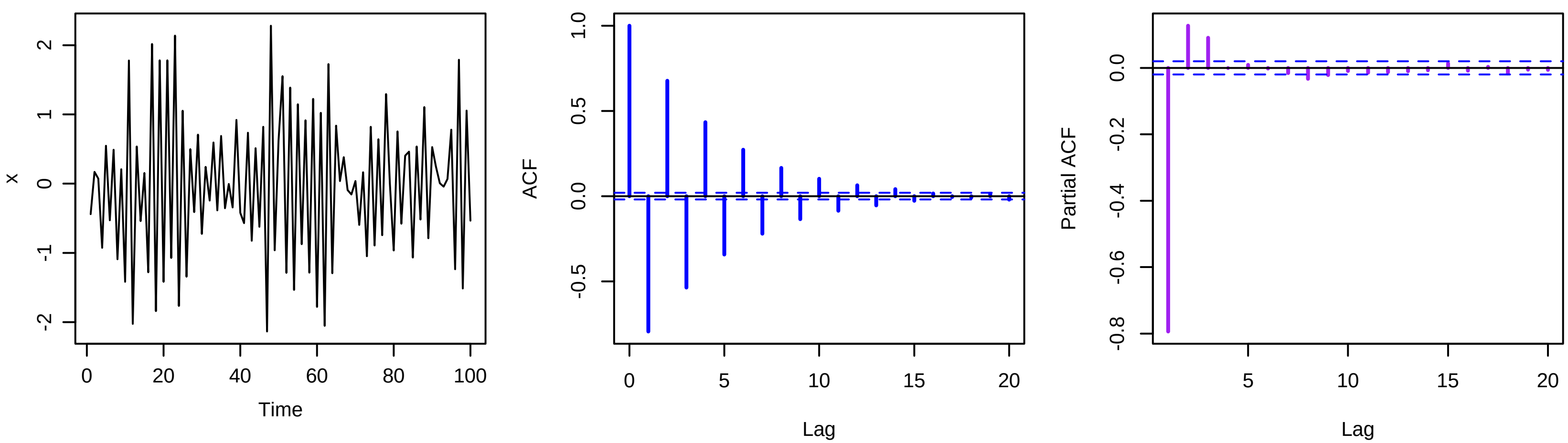
$$x_0^{k-1} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{k-1} x_{k-1}$$

ACF & PACF for AR(3) processes

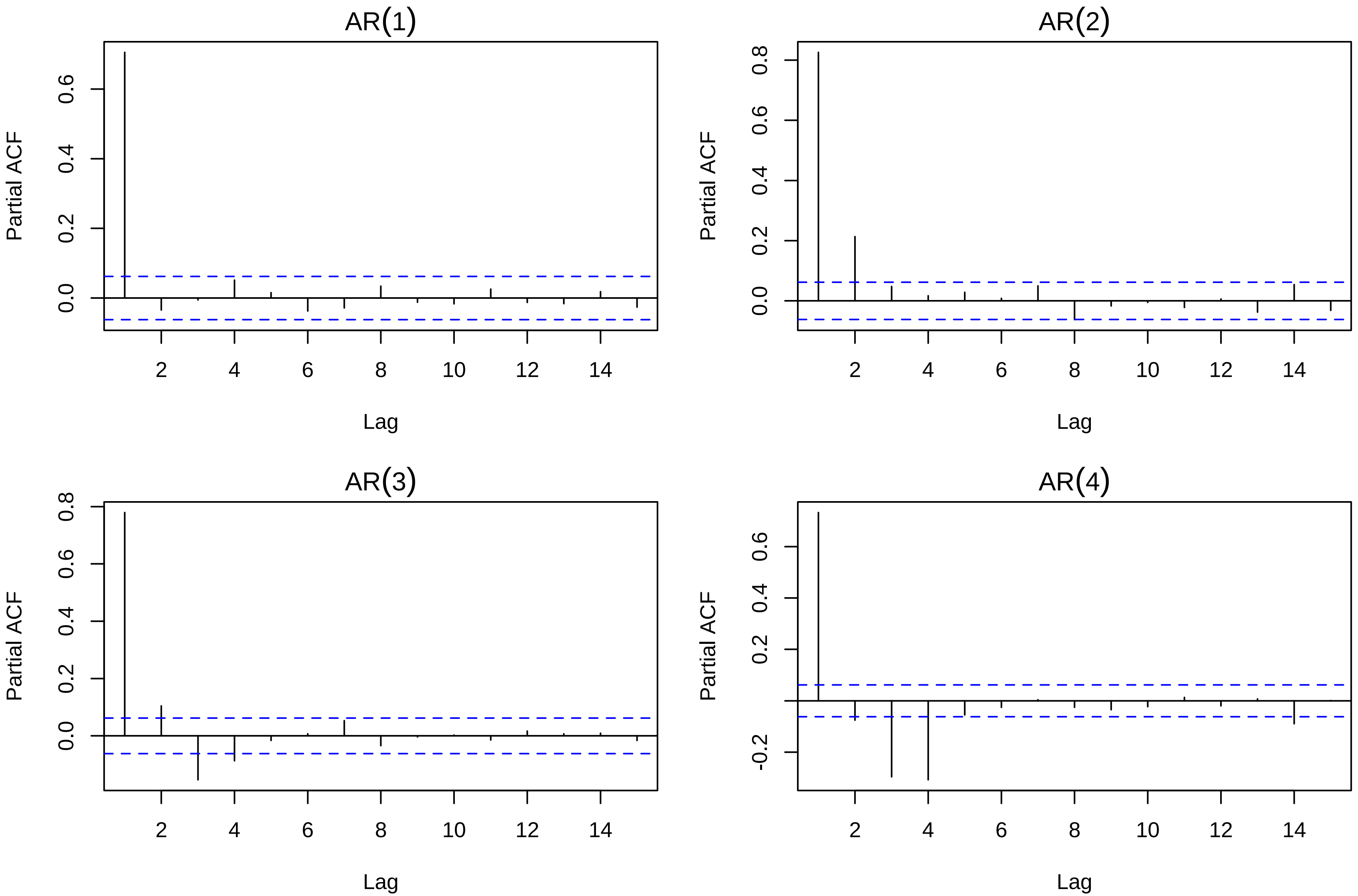
AR(3) with $f_1 = 0.7$, $f_2 = 0.2$, $f_3 = -0.1$



AR(3) with $f_1 = -0.7$, $f_2 = 0.2$, $f_3 = -0.1$



PACF for AR(*p*) processes



Moving average (MA) models

- A *moving average* model of order *q*, or MA(*q*), is defined as

$$x_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}$$

where w_t is WN (with 0 mean)

- It is simply the current error term plus a weighted sum of the *q* most recent error terms

- Because MA processes are finite sums of stationary WN processes, they are themselves stationary

Invertible MA models

- We can write out an MA(q) model using the backward shift notation, such that

$$x_t = \left(1 + \theta_1 \mathbf{B} + \theta \mathbf{B}^2 + \ldots + \theta_q \mathbf{B}^q\right) w_t = \theta_q\left(\mathbf{B}\right) w_t$$

- An MA process is *invertible* if it can be expressed as a stationary autoregressive process of infinite order without an error term

- For example, an MA(1) process

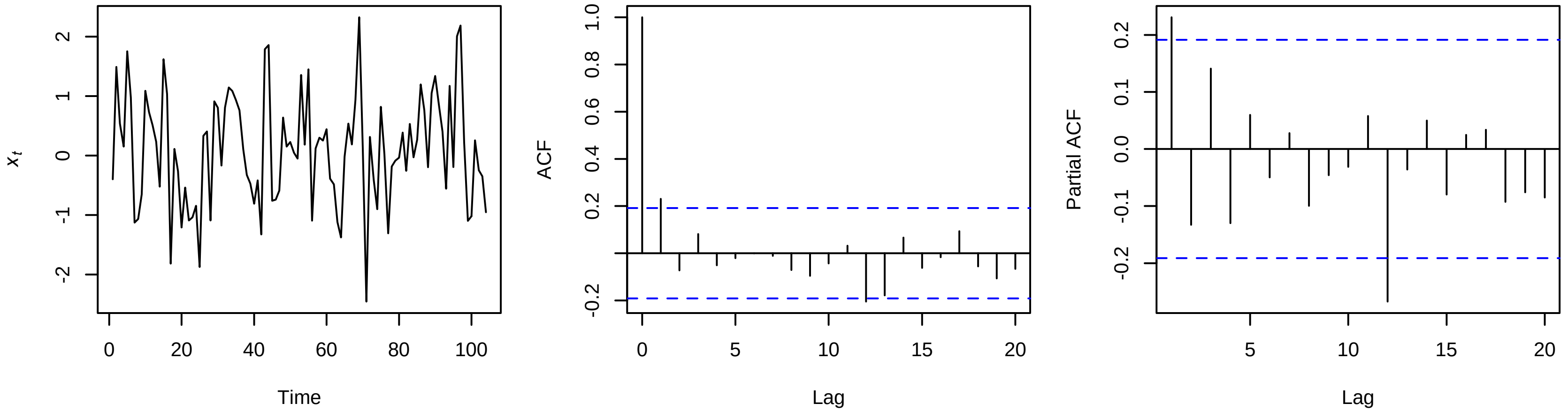
$$x_t = \left(1 - \theta \mathbf{B}\right) w_t$$

$$w_t = \left(1 - \theta \mathbf{B}\right)^{-1} x_t$$

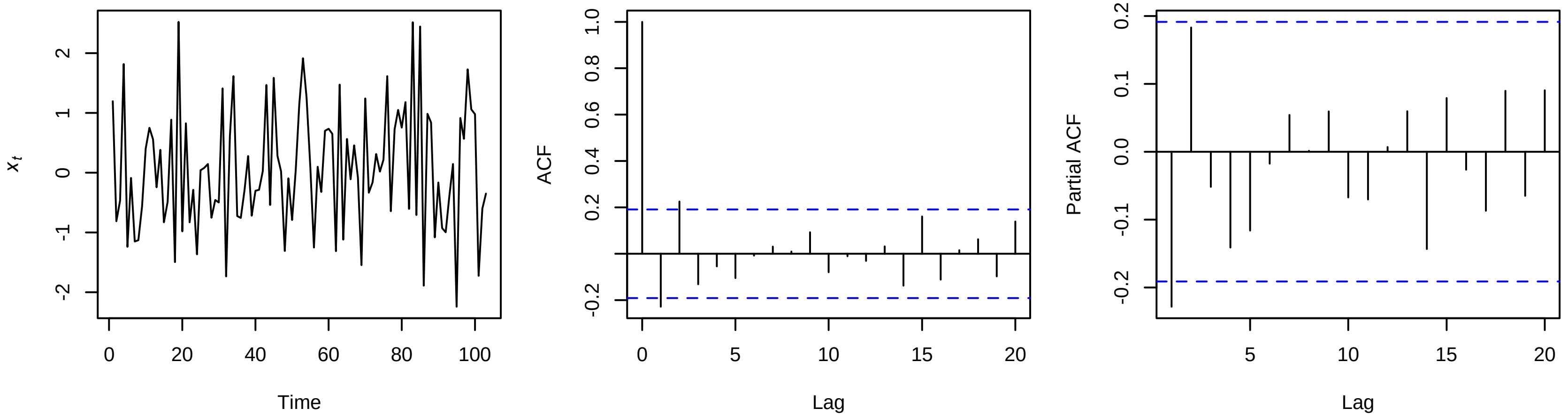
$$w_t = \left(1 + \theta \mathbf{B} + \theta^2 \mathbf{B}^2 + \ldots\right) x_t = x_t + \theta x_{t-1} + \theta^2 x_{t-2} + \ldots$$

Examples of MA(q) processes

MA(1) with q = 0.3



MA(2) with q_1 = -0.3, q_2 = 0.3



Autoregressive moving average models

- A time series is *autoregressive moving average*, or ARMA(*p*,*q*), if it is stationary and

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

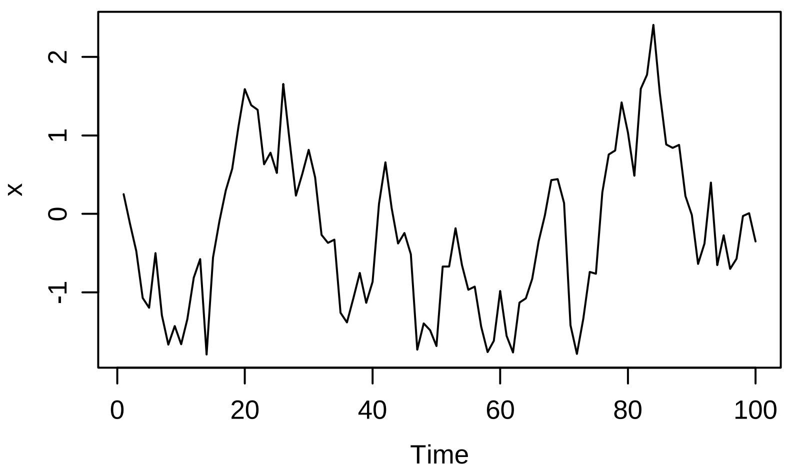
- We can write out an ARMA(*p*,*q*) model using the backward shift notation, such that

$$\phi_p(\mathbf{B})x_t = \theta_q(\mathbf{B})w_t$$

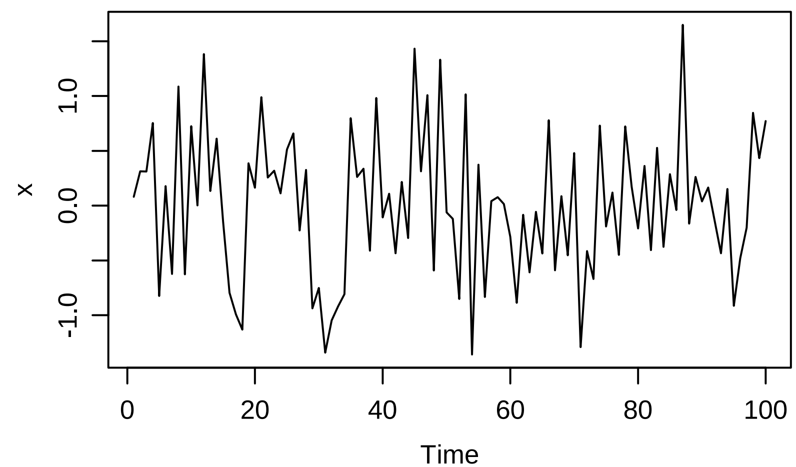
- ARMA models are *stationary* if all roots of $\phi > 1$
- ARMA models are *invertible* if all roots of $\theta > 1$

Examples of ARMA(*p*,*q*) processes

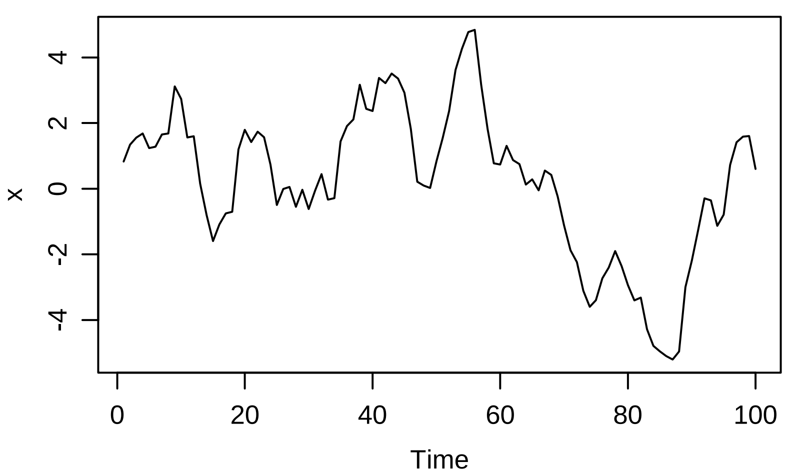
ARMA(3,1): $f_1 = 0.7$, $f_2 = 0.2$, $f_3 = -0.1$, $q_1 = 0.5$



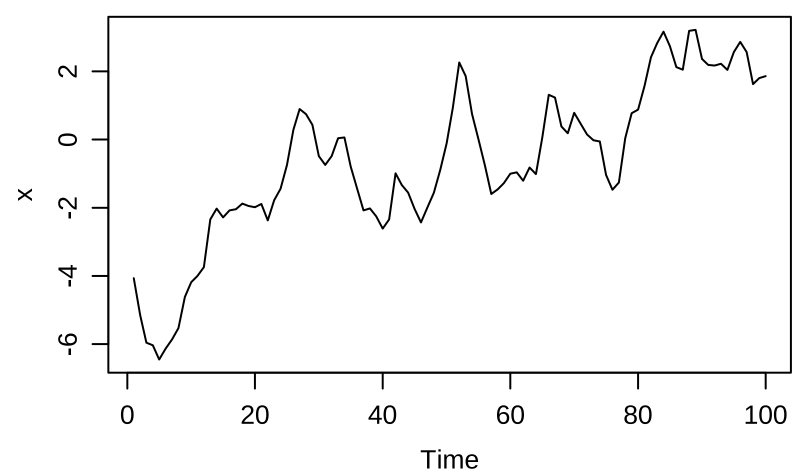
ARMA(2,2): $f_1 = -0.7$, $f_2 = 0.2$, $q_1 = 0.7$, $q_2 = 0.2$



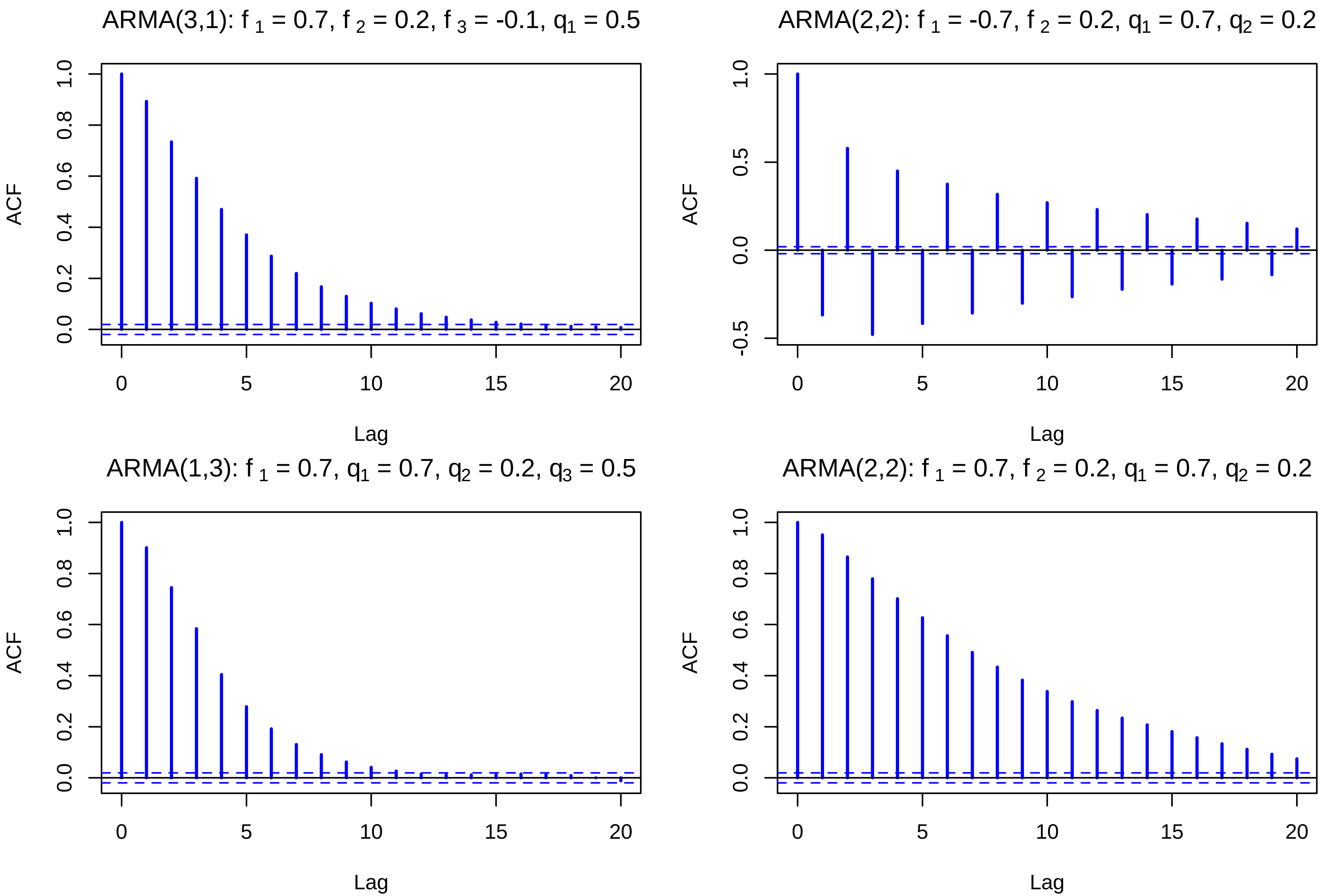
ARMA(1,3): $f_1 = 0.7$, $q_1 = 0.7$, $q_2 = 0.2$, $q_3 = 0.5$



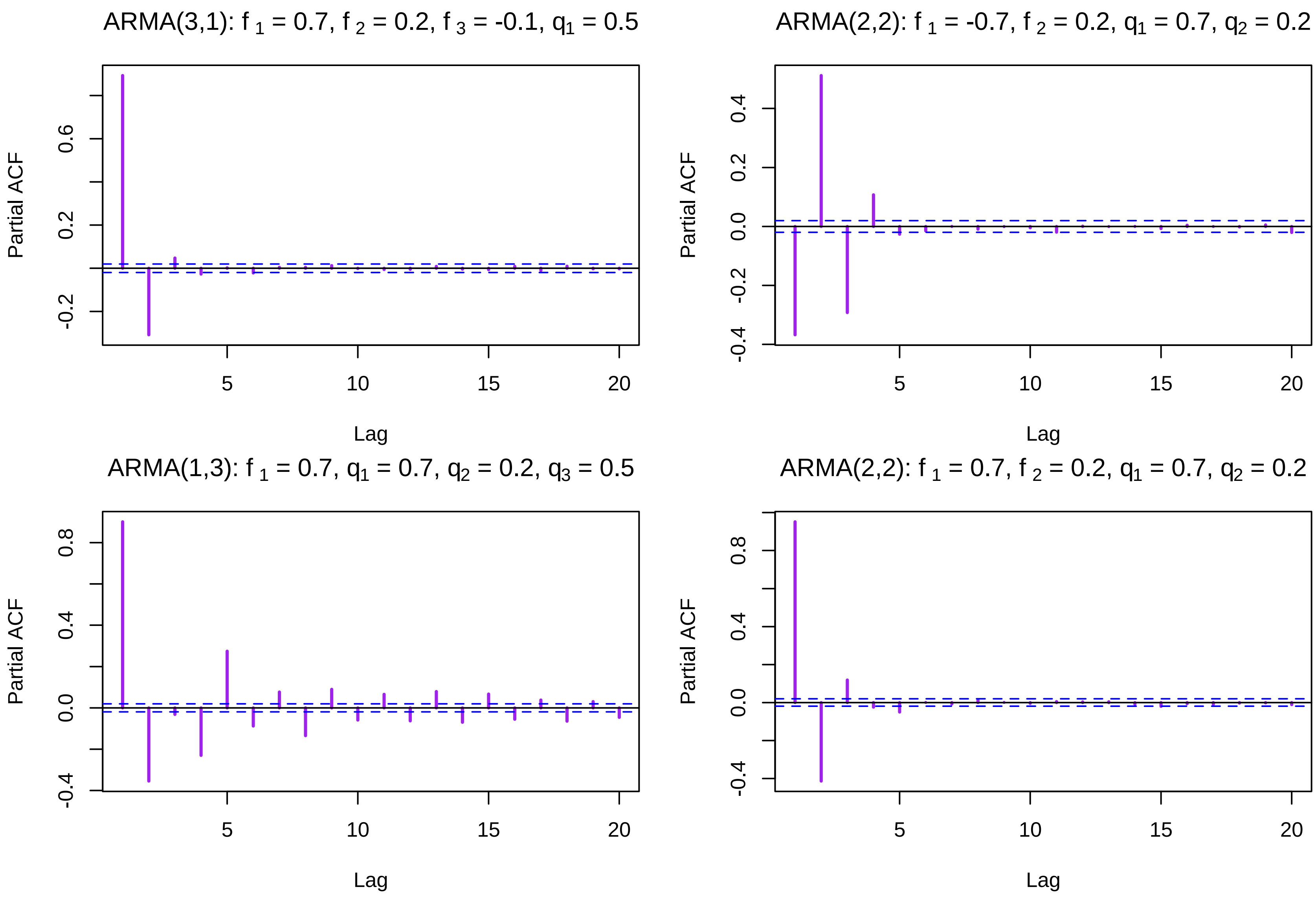
ARMA(2,2): $f_1 = 0.7$, $f_2 = 0.2$, $q_1 = 0.7$, $q_2 = 0.2$



ACF for ARMA(*p*,*q*) processes



PACF for ARMA(*p*,*q*) processes



Using ACF & PACF for model ID

	ACF	PACF
AR(p)	Tails off	Cuts off after lag-p
MA(q)	Cuts off after lag-q	Tails off
ARMA(p,q)	Tails off (after lag [q-p])	Tails off (after lag [p-q])

Topics for this lab

- `ts` class in R
- Plotting ts objects
- Understand covariance & correlation
- Examine some simple ts models
- Use `diff()` for trend/season removal
- Examine properties via `acf()` & `pacf()`
- Examine AR(p) models
- Examine MA(q) models
- ARMA(p,q) models via ‘arima.sim()’

Linear filtering of time series

- Beginning with the trend (m_t), we need a means for extracting a "signal"

- A common method is to use linear filters

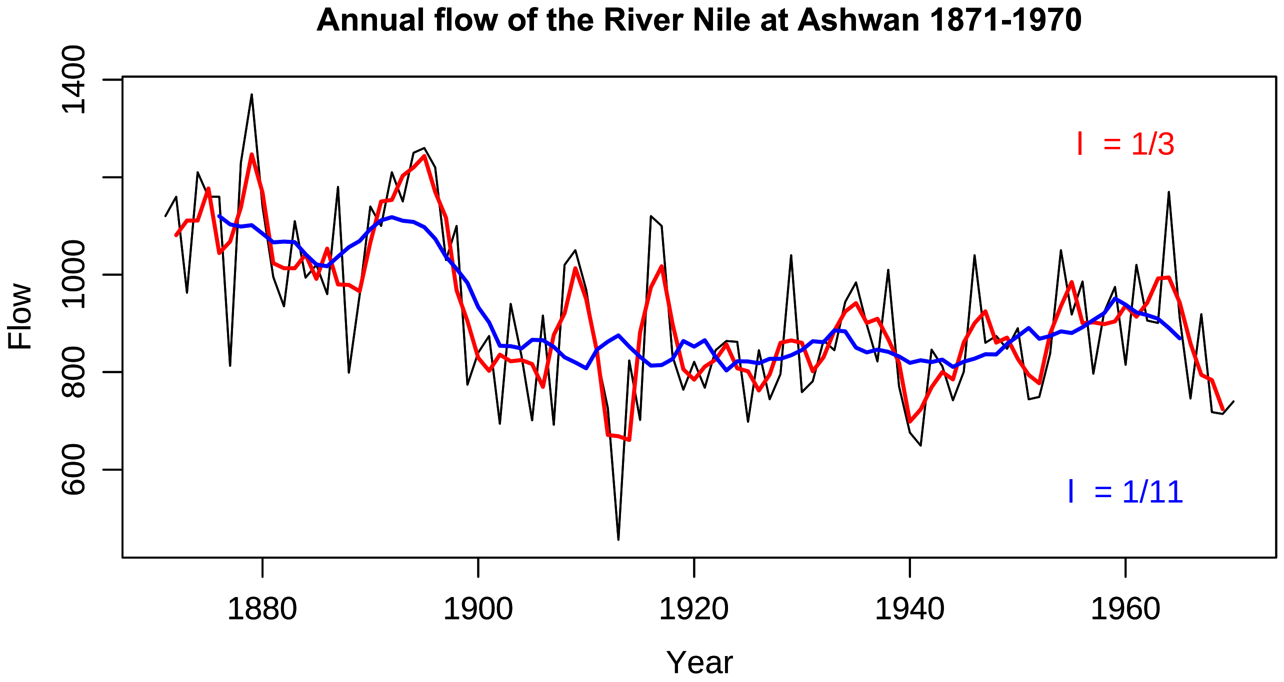
$$m_t = \sum_{i=-\infty}^{\infty} \lambda_i x_{t+i}$$

- For example, moving averages with equal weights

$$m_t = \sum_{i=-a}^{a} \frac{1}{2a+1} x_{t+i}$$

(FYI, this is what Excel does)

Example of linear filtering



Linear filtering of time series

- Consider case where season is based on 12 months & ts begins in January (t=1)
- Monthly averages over year will result in t = 6.5 for m_t (which is not good)
- One trick is to average (1) the average of Jan-Dec & (2) the average of Feb-Jan

$$m_t = \frac{\frac{1}{2}x_{t-6} + x_{t-5} + \cdots + x_{t-1} + x_t + x_{t+1} + \cdots + x_{t+5} + \frac{1}{2}x_{t+6}}{12}$$

Example of linear filtering